Multi-Carrier Multiple Access is Sum-Rate Optimal for Block Transmissions over Circulant ISI Channels*

Shuichi Ohno[1], Paul A. Anghel[2], Georgios B. Giannakis[2], and Zhi-Quan Luo[3]

[1] Dept. of Mathematics and Computer Science, Shimane Univ., Shimane, 690-8504, Japan

[2] Dept. of ECE, Univ. of Minnesota, 200 Union St. SE, Minneapolis, MN 55455, USA

[3] Dept. of ECE, McMaster Univ., Hamilton, Ontario, L8S 4K1, Canada

***Abstract*— We establish that practical multiple access based on `nite size information blocks transmitted with prescribed power and with loaded multicarrier modulation, is optimal with respect to maximizing the sum-rate of circulant inter-symbol interference (ISI) channels, that are assumed available at the transmitter. Circulant ISI channels are ensured either with cyclic pre`xed block transmissions and an overlap-save reception, or, with zero-padded block transmissions and an overlap-add reception. Analysis asserts that sum-rate optimal multicarrier users could share one or more subcarriers depending on the underlying channels. Optimal loading is performed by specializing an existing iterative low-complexity algorithm to circulant ISI channels.**

I. Introduction

It has been long recognized that both performance and data rates of transmissions over ISI channels can be optimized, when Channel State Information (CSI) is made available at the transmitter; e.g., via feedback, or, during a time-division duplex session. Single-user multi-carrier transmissions loaded according to the CSI-based "water-pouring" principle are known to achieve the ISI channel capacity for a prescribed power budget, provided that in`nitely many, and in`nitely long complex exponentials are utilized as information bearing waveforms [6].

Such ideal transmissions are serial, and assume in`nitely long information blocks. On the other hand, power- and bit-loaded multicarrier transmissions with `nite size blocks implement water-pouring in practice. The resulting Discrete Multi-Tone (DMT) systems, have been standardized for communication over digital subscriber lines (x-DSL) [7]. Thanks to the cyclic pre`x (CP) that is inserted per transmitted block, and is discarded from each received block, DMT systems enable block-by-block processing without ISI-induced inter-block interference (IBI), as in an overlap-save implementation of block convolution. Equally important, the CP insertion and removal renders the underlying channel matrix (in the discrete-time baseband equivalent block DMT model) circulant. It has been shown that DMT maximizes the capacity of the resulting circulant ISI channel [7] (see also [9] for generalizations to linear block convolution models entailing Toeplitz channel matrices).

Interestingly, similar optimality with respect to sum-capacity has not been fully established for practical *multiple access based on finite size blocks* transmitted through multiuser ISI channels. However, conditions for maximizing the sum-capacity of multiuser ISI channels have revealed that frequency-division multiple access (FDMA) offers an optimal solution in the ideal case, where each user s serial transmissions are presumably `ltered with in`nitely long `lters to become Gaussian, and are loaded according to a multiuser water-pouring principle [2]. But it was not until recently, that a practical (albeit suboptimum) algorithm was devised in [12] to maximize the conditions in [2]. Interesting related algorithms were presented earlier in [4], and were purported optimal by restricting the class of multiple access schemes to the practically attractive DMT class; i.e., along with the aforementioned CP insertion and removal, it was *a fortiori* assumed in [4] that each user relies on multi-carrier modulation.

In this paper, we prove that multi-carrier multiple access is sum-rate optimal for `nite block `xed-power transmissions over circulant ISI channels. One case that circulant ISI channels arise is when each user transmits blocks with a CP that is discarded at the receiver, as in the single user DMT. Another case is when each user pads zeros per transmitted block (at least as many as the ISI channel order), and an overlap-add operation is performed per received block, as detailed in [10]. Unlike [4], where DMT per user is assumed, we prove here that judiciously loaded DMT per user possesses sum-rate optimality. Implementation of the optimal loading follows the iterative algorithm of [11], which was developed for general channel matrices, and shown to have low complexity (linear in the number of users and the block length). It turns out that applying [11] to circulant channel matrices, leads to a power loading algorithm that is less complex than the one in [4]. Although in the limit our `nite block optimal multi-carrier transmissions should coincide with the asymptotically optimal FDMA scheme in [2], it is shown possible with `nite blocks to have users sharing subcarriers. The same conclusion was also reached by the multiple access scheme in [8], which relies on combinatorial or semi-de`nite programming algorithms to optimize the sum-mean-square symbol error criterion, under transmit-power constraints per user.

We adopt the following notational conventions: bold upper (lower) case denotes matrices (column vectors). An upper case subscript associated with a matrix denotes its dimension. An $N \times N$ identity matrix is denoted with $\boldsymbol{I}_N$, and the all-zero matrix of size $N \times L$ is denoted with $\mathbf{0}_{N\times L}$. The determinant of a matrix $\boldsymbol{A}$ is denoted by $|\boldsymbol{A}|$. Superscript $\mathcal{H}$ will stand for Hermitian and ^{T} denotes transpose. We take $[x]_+$ to mean that $[x]_+ = x$ if $x > 0$, otherwise $[x]_+ = 0$. We use $\mathrm{diag}(\boldsymbol{x})$ to denote a diagonal matrix with $\boldsymbol{x}$ on its diagonal. We denote by $\boldsymbol{F}$ the $N \times N$ FFT matrix with the $(i+1, k+1)$th entry $\exp(-j2\pi ik/N)/\sqrt{N}$. Ensemble averaging is denoted with $\mathrm{E}\{\cdot\}$.

II. Preliminaries

We consider block transmissions for multiple access over frequency selective ISI channels. Fig. 1 shows the discrete-time baseband equivalent uplink communication model of the mth

*Work in this paper was supported by the NSF Wireless Initiative grant no. 99-79443

0-7803-7400-2/02/$17.00 © 2002 IEEE

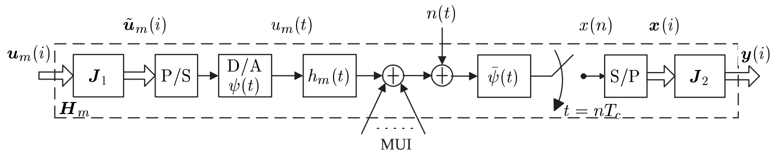

Fig. 1. Continuous and discrete-time baseband equivalent multichannel model.

user, where $m \in [1, M]$, and M is the maximum number of users in the system. Each $N \times 1$ block of data symbols $\boldsymbol{u}_m(i)$ has zero-mean, and its power is upper bounded:

$$\mathrm{E}\{||\boldsymbol{u}_m(i)||^2\} \leq \mathcal{P}_m, \qquad m \in [1, M]. \tag{1}$$

Before transmission, $\boldsymbol{u}_m(i)$ is multiplied by a tall $(N+L) \times N$ matrix $\boldsymbol{J}_1 \triangleq [[\mathbf{0}_{L\times(N-L)}, \boldsymbol{I}_L]^T, \boldsymbol{I}_N]^T$, which inserts a CP of length $L < N$. The mth user's $(N+L) \times 1$ cyclic-prefixed block $\tilde{\boldsymbol{u}}_m(i) \triangleq \boldsymbol{J}_1 \boldsymbol{u}_m(i)$ is parallel-to-serial (P/S) converted to a sequence $\tilde{u}_m(n)$, which is subsequently pulse-shaped to yield the continuous-time signal $u_m(t) \triangleq \sum_{n=-\infty}^{+\infty} \tilde{u}_m(n)\psi(t-nT_c)$, where $\psi(t)$ is the chip-pulse, and T_c denotes the chip period. The mth user's transmitted signal $u_m(t)$ propagates through a channel $h_m(t)$. We assume that:
a1) the channels $\{h_m(t)\}_{m=1}^M$ are known at the mobile transmitters, and have delay spreads upper bounded by LT_c.
The received signal $x(t)$ is filtered by a square-root Nyquist receive-filter $\bar{\psi}(t)$, and then sampled at the chip rate $1/T_c$ to yield $x(n)$. Let us denote with $h_m(l) \triangleq (\psi \star h_m \star \bar{\psi})(t)|_{t=lT_c}$ the mth user's equivalent chip-sampled channel impulse response. As per a1), the maximum order of $\{h_m(l)\}_{m=1}^M$ is L. In addition to transmit-receive filters $\{\psi(t), \bar{\psi}(t)\}$ and multipath effects, the FIR channel $h_m(l)$ captures the mth user's asynchronism in the form of delay factors. With mobile users attempting to synchronize with the base station's pilot waveform, the relative asynchronism among users is supposed to be only a few chips in duration (quasi-synchronous or block-synchronous setup).

The received sequence $x(n)$ is serial-to-parallel (S/P) converted to form blocks $\boldsymbol{x}(i) \triangleq [x(i(N+L)), x(i(N+L)+1), \ldots, x(i(N+L)+N+L-1)]^{\mathrm{T}}$ of length $N+L$. We remove the CP from $\boldsymbol{x}(i)$ using the matrix $\boldsymbol{J}_2 \triangleq [\mathbf{0}_{N\times L}, \boldsymbol{I}_N]$, and express the resulting IBI-free $N \times 1$ received block as:

$$\boldsymbol{y}(i) = \boldsymbol{J}_2 \boldsymbol{x}(i) = \sum_{m=1}^{M} \boldsymbol{H}_m \boldsymbol{u}_m(i) + \boldsymbol{v}(i), \tag{2}$$

where $\boldsymbol{H}_m$ is an $N \times N$ circulant matrix with (n,k)th entry $h_m((n-k) \bmod N)$, and $\boldsymbol{v}(i)$ is an $N \times 1$ complex Gaussian noise vector with zero mean, and correlation matrix $\boldsymbol{R}_v \triangleq \boldsymbol{D}_v^{-1} = \mathrm{diag}(\rho_0^{-1}, \ldots, \rho_{N-1}^{-1})$. Note that white noise is not colored after being processed with $\boldsymbol{J}_2$. Consequently, if we assume that white noise $\boldsymbol{\eta}(i)$ is present in $\boldsymbol{x}(i)$, then $\boldsymbol{v}(i) \triangleq \boldsymbol{J}_2\boldsymbol{\eta}(i)$ is also white and $\boldsymbol{R}_v = \rho^{-1}\boldsymbol{I}_N$.

We allow for non-identical noise variances to encompass block transmissions with zero-padding (ZP) replacing the CP, where $\tilde{\boldsymbol{u}}_m(i)$ is replaced by $\bar{\boldsymbol{u}}_m(i) \triangleq [\boldsymbol{I}_N^T, \mathbf{0}_{N\times L}^T]^T \boldsymbol{u}_m(i)$. This avoids IBI, and leads to circulant channel matrices too, provided that an overlap-add operation is performed on each received block $\boldsymbol{x}(i)$ to yield $\boldsymbol{y}(i) \triangleq [\boldsymbol{I}_N, \boldsymbol{I}_{zp}]\boldsymbol{x}(i)$, where $\boldsymbol{I}_{zp}$ denotes the first L columns of $\boldsymbol{I}_N$ [10]. The main difference with the CP-based model (2) is that white noise is colored by the overlap-add operation, and its covariance matrix becomes diagonal with unequal diagonal entries.

We further assume that:
a2) the transmitted blocks $\{\boldsymbol{u}_m(i)\}_{m=1}^M$ are zero mean, mutually uncorrelated with covariance matrices $\{\boldsymbol{R}_{u_m}\}_{m=1}^M$, and also uncorrelated with the noise $\boldsymbol{v}(i)$.
For ease of notation, we will subsequently drop the block index i, since it is common to all vectors in (2).

The maximum sum of achievable rates for the users in the system can be computed as the maximum mutual information between the transmitted blocks $\{\boldsymbol{u}_m\}_{m=1}^M$, and the received block $\boldsymbol{y}$, denoted by $I(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M; \boldsymbol{y})$, over the joint pdf $p(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M)$:

$$\mathcal{C}_S = \max_{p(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M)} I(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M; \boldsymbol{y}). \tag{3}$$

For the multi-access channel defined by the input/output relation (2), the sum-capacity is achieved if and only if $\boldsymbol{y}$ is Gaussian (e.g., [3, p. 254]). With $\{\boldsymbol{u}_m\}_{m=1}^M$ jointly Gaussian, $\boldsymbol{y}$ is Gaussian too, and (3) can then be rewritten as [c.f. a2)]

$$\mathcal{C}_S = \max_{\{\boldsymbol{R}_{u_m}\}_{m=1}^M} \log \left| \boldsymbol{I}_N + \boldsymbol{D}_v \sum_{m=1}^{M} \boldsymbol{H}_m \boldsymbol{R}_{u_m} \boldsymbol{H}_m^{\mathcal{H}} \right|. \tag{4}$$

Given a1), the objective of this paper is to specify the correlation matrices $\{\boldsymbol{R}_{u_m}\}_{m=1}^M$ that achieve the sum-capacity in (4), subject to the power constraint of (1), which we write as

$$\mathrm{trace}\{\boldsymbol{R}_{u_m}\} \leq \mathcal{P}_m, \qquad m \in [1, M]. \tag{5}$$

Because $\boldsymbol{R}_{u_m}$ is Hermitian, we only have to specify the $N(N+1)/2$ elements on, and above its main diagonal. Hence, the maximization in (4) entails $MN(N+1)/2$ variables.

It is important to point out that we have not assumed a priori that each user adopts a multicarrier modulation, like in the multiuser DMT system of [4]. Even though we do not assume FFT processing in (2), we show in the next theorem that for block transmissions over circulant ISI channels, a DMT type of precoder with appropriately loaded subcarriers is optimal in the sense of maximizing (4) subject to the constraint (5).

III. OPTIMALITY OF MULTICARRIER MULTI-ACCESS

We first state our main result:

Theorem 1 *For multi-access through circulant ISI channels, the sum-capacity of the MIMO system described by* (2) *is achieved if the transmitted blocks are zero-mean Gaussian with correlation matrix*

$$\boldsymbol{R}_{u_m} = \boldsymbol{F}^{\mathcal{H}} \boldsymbol{\Lambda}_m \boldsymbol{F}, \qquad \forall m \in [1, M], \tag{6}$$

where $\boldsymbol{\Lambda}_m \triangleq \mathrm{diag}(\lambda_{m,0}, \lambda_{m,1}, \ldots, \lambda_{m,N-1})$.

Proof: The circulant channel matrix $\boldsymbol{H}_m$ can be diagonalized by multiplying it to the left and right with the FFT matrix $\boldsymbol{F}$ and the IFFT matrix $\boldsymbol{F}^{\mathcal{H}}$, respectively, to obtain

$$\boldsymbol{F}\boldsymbol{H}_m\boldsymbol{F}^{\mathcal{H}} = \mathrm{diag}(H_m(\omega_0), \ldots, H_m(\omega_{N-1})) \triangleq \boldsymbol{D}_{H_m}, \tag{7}$$

where $H_m(\omega_n)$ is the transfer function of the mth user's channel $h_m(l)$ at frequency $\omega_n = 2\pi n/N$. We substitute $\boldsymbol{H}_m$ from (7) into (4) to find

$$\mathcal{C}_S = \max_{\{\boldsymbol{R}_m\}_{m=1}^M} \log\left|\boldsymbol{F}^{\mathcal{H}}(\boldsymbol{I}_N+\boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\boldsymbol{F}\boldsymbol{R}_{u_m}\boldsymbol{F}^{\mathcal{H}}\boldsymbol{D}_{H_m}^{\mathcal{H}})\boldsymbol{F}\right|$$
$$= \max_{\{\boldsymbol{R}_m\}_{m=1}^M} \log\left|\boldsymbol{I}_N+\boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}\right|,$$

where $\tilde{\boldsymbol{R}}_m \triangleq \boldsymbol{F}\boldsymbol{R}_{u_m}\boldsymbol{F}^{\mathcal{H}}$. Since $\mathrm{trace}\{\boldsymbol{R}_{u_m}\} = \mathrm{trace}\{\tilde{\boldsymbol{R}}_m\}$, our optimization problem is equivalent to:

$$\max_{\{\tilde{\boldsymbol{R}}_m\}_{m=1}^M} \log\left|\boldsymbol{I}_N+\boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}\right| \quad (8)$$

subject to

$$\mathrm{trace}\{\tilde{\boldsymbol{R}}_m\} \le \mathcal{P}_m, \quad m \in [1, M]. \quad (9)$$

Matrix $\boldsymbol{I}_N + \boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}$ is positive definite, and consequently, we can apply the Hadamard inequality [3, p. 502] to $|\boldsymbol{I}_N + \boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}|$ to find that

$$\log\left|\boldsymbol{I}_N+\boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}\right| \quad (10)$$
$$\le \log\left[\prod_{n=0}^{N-1}\left(1+\rho_n\sum_{m=1}^{M}|H_m(\omega_n)|^2[\tilde{\boldsymbol{R}}_m]_{n,n}\right)\right] \quad (11)$$
$$= \sum_{n=0}^{N-1}\log\left(1+\rho_n\sum_{m=1}^{M}|H_m(\omega_n)|^2[\tilde{\boldsymbol{R}}_m]_{n,n}\right), \quad (12)$$

where $[\tilde{\boldsymbol{R}}_m]_{n,n}$ denotes the (n,n)th entry of $\tilde{\boldsymbol{R}}_m$, and the equality holds if and only if $\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\tilde{\boldsymbol{R}}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}$ is diagonal. We seek correlation matrices $\tilde{\boldsymbol{R}}_m$ maximizing (10) under the power constraint in (9). Since (9) and (11) *do not depend on the off-diagonal entries* of the correlation matrices, we can set them to zero in order to obtain a diagonal $\tilde{\boldsymbol{R}}_m$ maximizing (10). With these diagonal $\tilde{\boldsymbol{R}}_m$, the equality holds in (11). This implies that the sum-capacity is attained when $\tilde{\boldsymbol{R}}_m$ are diagonal, call them $\boldsymbol{\Lambda}_m$; hence, $\boldsymbol{R}_{u_m}$ can be expressed as $\boldsymbol{R}_{u_m} = \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\Lambda}_m\boldsymbol{F}$, which completes the proof. ■

We remark that unlike the single user water-pouring setup, multicarrier modulation offers only sufficient sum-rate maximizing transmissions for multiple access through ISI channels: indeed, having the sum of matrices in (10) to be diagonal is necessary and sufficient for sum-rate optimality; however, having every summand matrix to be diagonal is only sufficient.

Let us denote the block of information symbols of user m with $\boldsymbol{s}_m$, and with $\boldsymbol{R}_{s_m} \triangleq \mathrm{E}\{\boldsymbol{s}_m\boldsymbol{s}_m^{\mathcal{H}}\}$ its correlation matrix. Notice that colored $\boldsymbol{s}_m$'s could result either from coding, or, from linear precoding [10]. In both cases it is possible to have $\boldsymbol{s}_m$ inherit the optimal transmit-correlation matrix (6) via a linear transformation $\boldsymbol{\Theta}_m$. Indeed, using

$$\boldsymbol{\Theta}_m = \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\Lambda}_m^{\frac{1}{2}}\boldsymbol{R}_{s_m}^{-\frac{1}{2}}, \quad \boldsymbol{u}_m = \boldsymbol{\Theta}_m\boldsymbol{s}_m, \quad m \in [1, M], \quad (13)$$

we can verify that $\boldsymbol{R}_{u_m}$ satisfies (6). Notice that $\boldsymbol{R}_{s_m}^{-\frac{1}{2}}$ prewhitens the possibly correlated $\boldsymbol{s}_m$, while $\boldsymbol{\Lambda}_m^{\frac{1}{2}}$ appropriately loads the N symbols of user m, before placing them on the N digital subcarriers (columns of the N-point IFFT matrix $\boldsymbol{F}^{\mathcal{H}}$). We have thus established the main claim of this paper:

Corollary 1 *Loaded multicarrier modulation based on finite size information blocks transmitted with prescribed power is sum-rate optimal for transmissions over circulant ISI channels.*

To complete our sum-capacity optimization, we need to specify the optimum set $\{\boldsymbol{\Lambda}_m\}_{m=1}^M$ that maximizes

$$g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1}) \triangleq \log\left|\boldsymbol{I}_N+\boldsymbol{D}_v\sum_{m=1}^{M}\boldsymbol{D}_{H_m}\boldsymbol{\Lambda}_m\boldsymbol{D}_{H_m}^{\mathcal{H}}\right|$$
$$= \sum_{n=0}^{N-1}\log\left(1+\rho_n\sum_{m=1}^{M}|H_m(\omega_n)|^2\lambda_{m,n}\right), \quad (14)$$

with respect to the diagonal entries $\{\lambda_{m,n}\}_{n=0}^{N-1}$ of $\boldsymbol{\Lambda}_m$, and subject to

$$\mathrm{trace}\{\boldsymbol{\Lambda}_m\} = \sum_{n=0}^{N-1}\lambda_{m,n} \le \mathcal{P}_m, \quad m \in [1, M]. \quad (15)$$

Thanks to Theorem 1, the number of optimization variables has been reduced from $MN(N+1)/2$ in our initial maximization problem, to MN in (14).

In order to solve this maximization problem, we rely on the algorithm of [11], which has been developed for a multi-antenna multiple access system, but also applies to our case. Theorem 1 of [11] asserts that $\{\boldsymbol{R}_{u_m}\}_{m=1}^M$ solve the sum-rate maximization problem in (4), subject to the constraint (5), if and only if $\boldsymbol{R}_{u_m}$ is the single user water-filling covariance matrix for the channel $\boldsymbol{H}_m$, and for the noise plus multi-user interference (NMUI) covariance matrix $\boldsymbol{D}_v^{-1}+\sum_{\mu=1,\mu\neq m}^{M}\boldsymbol{H}_\mu\boldsymbol{R}_{u_\mu}\boldsymbol{H}_\mu^{\mathcal{H}}$, for all $m \in [1, M]$, simultaneously. This theorem leads to a simple and efficient water-filling algorithm in which each $\boldsymbol{R}_{u_m}$, $m \in [1, M]$, is computed iteratively after the NMUI has been updated following each computation of $\boldsymbol{R}_{u_m}$. However, similar to [9], each iterative calculation of $\boldsymbol{R}_{u_m}$ requires singular value decomposition (SVD).

Because for our setup the sum-capacity is achieved by $\boldsymbol{R}_{u_m}$'s with the structure of (6), we can specialize the algorithm of [11] so that at each iteration the algorithm just updates $\{\lambda_{m,n}\}_{n=0}^{N-1}$, and does not perform an SVD operation since given $\{\lambda_{\mu,n}\}_{n=0}^{N-1}$, $\mu \neq m$, the optimal $\{\lambda_{m,n}\}_{n=0}^{N-1}$ can be computed analytically as in [6, p. 334]. We now present our modified version of the algorithm in [11] that is tailored for circulant ISI channels:

Algorithm 1

initialize $\lambda_{m,n} = 0, \quad m \in [1, M], \; n \in [0, N-1]$
repeat
  for $m = 1$ to M

$$d_{m,n} = \rho_n \sum_{\mu=1,\mu\neq m}^{M}|H_\mu(\omega_n)|^2\lambda_{\mu,n}, \quad \forall n \in [0, N-1]$$

  for $n = 0$ to $N-1$
    compute $\lambda_{m,n}$ as the single user water-pouring solution :

$$\begin{cases}\lambda_{m,n} = \left[\alpha_m - \frac{1+d_{m,n}}{\rho_n|H_m(\omega_n)|^2}\right]_+, \\ \text{where } \alpha_m \text{ is chosen such that } \sum_{n=0}^{N-1}\lambda_{m,n} = \mathcal{P}_m.\end{cases}$$

  end
end

until variation of $g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1})$ is less than ε.

Algorithm 1 belongs to the class of optimization algorithms which use parallel variable distribution. It is known that if the constraint set is in the form of a Cartesian product (i.e., no coupling between the sets of variables of different users) then the algorithm converges to a stationary point [5]. From Theorem 1 in [11] this stationary point achieves the sum-capacity of circulant ISI channels.

Notice that there are three computational loops in Algorithm 1. The inner loop is used to compute the optimal $\lambda_{m,n},\ \forall n \in [0, N-1]$ given all $\lambda_{\mu,n}$ with $\mu \neq m$. The solution is computed analytically using Lagrange multipliers. Consequently, for a given user m, it takes N steps to update all $\lambda_{m,n}$ s. A second loop is used to compute $\lambda_{m,n}$ for all M users. The outmost loop iterates over the second loop until the desired accuracy is reached. Let us denote with N_I the number of times we run the outmost loop. Hence, this algorithm takes $\mathcal{O}(N_I M N)$ computational steps. Compared to [4], where the complexity is exponential in the number of users, the complexity here is linear.

IV. Characterization of the Optimal Solution

In this subsection we establish similar properties for our optimal solutions, which we denote with $\{\lambda^*_{m,n}\}_{m=1,n=0}^{M,N-1}$, as the ones reported in [8], where the total mean square error of a multi-access system with ISI channels is minimized. Similar to [8], we partition the integer index set $I \triangleq \{0, 1, \ldots, N-1\}$ as $I = (\bigcup_{m=1}^{M} I_m) \bigcup I_{share} \bigcup I_{null}$, where

$$I_m = \{n | \lambda^*_{m,n} > 0, \lambda_{\mu,n} = 0, \mu \neq m,\ \mu \in [1, M]\},\ m \in [1, M],$$

$$I_{share} = \{n | \lambda^*_{m,n} > 0, \lambda^*_{\mu,n} > 0 \text{ for } m \neq \mu,\ m, \mu \in [1, M]\},$$

$$I_{null} = \{n | \lambda^*_{m,n} = 0,\ \forall m \in [1, M]\},$$

with I_m denoting the set of subcarriers of user m, I_{share} the set of subcarriers shared by the users in the system, and I_{null} the set of subcarriers not used by any user. Now, we can state a theorem which gives insight into the appropriate allocation of subcarriers.

Theorem 2 *For the subcarrier partitioning it holds that:*

a) $\dfrac{|H_m(\omega_n)|^2}{|H_\mu(\omega_n)|^2} \geq \dfrac{|H_m(\omega_p)|^2}{|H_\mu(\omega_p)|^2}$ *for any* $n \in I_m$ *and* $p \in I_\mu$.

b) $\dfrac{|H_m(\omega_n)|^2}{|H_\mu(\omega_n)|^2} = \dfrac{|H_m(\omega_p)|^2}{|H_\mu(\omega_p)|^2}$ $\forall n, p$ *shared by users* $m,\ \mu$.

c) $|H_m(\omega_n)|^2 < |H_m(\omega_p)|^2$ *for any* $n \in I_{null}$, *and* $p \in I_m \bigcup I_{share},\ \forall m \in [1, M]$.

Proof: The function $g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1})$ in (14) is concave because its domain is a convex set and the Hessian of $-g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1})$ is positive semi-definite. Since $g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1})$ is concave and the constraints in (15) are linear, we can apply Kuhn-Tucker s theorem [1, p. 249], which asserts that there exist Lagrange multipliers $\nu_m \geq 0$ for $m \in [1, M]$ satisfying $\forall n \in [0, N-1]$

$$\frac{|H_m(\omega_n)|^2}{1 + \rho_n \sum_{k=1}^{M} |H_k(\omega_n)|^2 \lambda^*_{k,n}} \leq \nu_m, \qquad \text{and}$$
$$\lambda^*_{m,n}\left(\frac{|H_m(\omega_n)|^2}{1 + \rho_n \sum_{k=1}^{M} |H_k(\omega_n)|^2 \lambda^*_{k,n}} - \nu_m\right) = 0. \tag{16}$$

For $n \in I_m$, we have $\lambda^*_{m,n} > 0$ and $\lambda^*_{\mu,n} = 0$ for $\mu \neq m$, so we can write (16) as

$$\frac{|H_m(\omega_n)|^2}{D_{m,n}} = \nu_m, \quad \frac{|H_\mu(\omega_n)|^2}{D_{m,n}} \leq \nu_\mu, \text{ for } \mu \neq m, \tag{17}$$

where $D_{m,n} \triangleq 1 + \rho_n |H_m(\omega_n)|^2 \lambda^*_{m,n}$. It follows from (17) that

$$\frac{|H_\mu(\omega_n)|^2}{|H_m(\omega_n)|^2} \leq \frac{\nu_\mu}{\nu_m},\ \forall n \in I_m. \tag{18}$$

Similar to (18), we obtain

$$\frac{|H_m(\omega_p)|^2}{|H_\mu(\omega_p)|^2} \leq \frac{\nu_m}{\nu_\mu},\ \forall p \in I_\mu. \tag{19}$$

By combining the inequalities (18) and (19), we obtain part a) of Theorem 2.

If $n \in I_{share}$, there exist at least two users m and μ with $\lambda^*_{m,n} > 0$ and $\lambda^*_{\mu,n} > 0$. We have from the second equation in (16) that

$$\frac{|H_m(\omega_n)|^2}{1 + \rho_n \sum_{k=1}^{M} |H_k(\omega_n)|^2 \lambda^*_{k,n}} = \nu_m$$
$$\frac{|H_\mu(\omega_n)|^2}{1 + \rho_n \sum_{k=1}^{M} |H_k(\omega_n)|^2 \lambda^*_{k,n}} = \nu_\mu, \tag{20}$$

which leads to

$$\frac{|H_m(\omega_n)|^2}{|H_\mu(\omega_n)|^2} = \frac{\nu_m}{\nu_\mu}. \tag{21}$$

Now, if subcarrier p is shared by the same users m and μ, similar to (21), we find

$$\frac{|H_m(\omega_p)|^2}{|H_\mu(\omega_p)|^2} = \frac{\nu_m}{\nu_\mu}. \tag{22}$$

By combining (21) and (22) we obtain part b) of the theorem.

Finally, if $n \in I_{null}$, then from (16) we have that $|H_m(\omega_n)|^2 \leq \nu_m$. For $p \in I_m \bigcup I_{share}$ we find from (17) and (20) that $\nu_m < |H_m(\omega_p)|^2$. It follows that $|H_m(\omega_n)|^2 < |H_m(\omega_p)|^2$, for any $n \in I_{null}$ and any $p \in I_m \bigcup I_{share}$, which proves part c) of our theorem. ■

Part a) of Theorem 2 asserts that allocation of subcarriers n and p between two users m and μ depends on the *relative* subchannel gains between the two users, namely, $|H_m(\omega_n)|^2/|H_\mu(\omega_n)|^2$ and $|H_m(\omega_p)|^2/|H_\mu(\omega_p)|^2$. More precisely, if we consider a two user system, similar to [8] the subcarriers for which $|H_1(\omega_n)|^2/|H_2(\omega_n)|^2$ is "high" are allocated to user 1, where as subcarriers for which $|H_1(\omega_n)|^2/|H_2(\omega_n)|^2$ is "low" are assigned to user 2.

From part b) of Theorem 2 we have that for all the subcarriers that are shared by two users, call them m and μ, the subchannel gain ratio $|H_m(\omega_n)|^2/|H_\mu(\omega_n)|^2$ is the same. Because our multi-access system is affected by fading, the channels $\{h_m(l)\}_{m=1}^{M}$ are realizations of some random channels. Consequently, the event for which $|H_m(\omega_n)|^2/|H_\mu(\omega_n)|^2$ is equal to $|H_m(\omega_p)|^2/|H_\mu(\omega_p)|^2$ for $n \neq p$ and for any $m, \mu \in [1, M]$, has measure zero. This implies that at most one subcarrier will be shared by any two users. Therefore, the total number of shared subcarriers is bounded by $M(M-1)/2$.

If we let the number of subcarriers $N \to \infty$, the superimposed spectra of the transmitted signals approaches an FDMA spectrum because the number of shared subcarriers is finite (a set of measure zero). This is in agreement with [2], which shows that FDMA achieves sum-capacity of the multi-access channel with ISI (note that adjacent user spectra in FDMA share a common zero frequency). However, when considering block transmissions with a `nite block size, the optimal spectrum can be achieved by loaded multicarrier transmissions, where the subcarriers from different users could eventually be shared.

Note, that for optimum transmissions, two users could share more than one subcarrier if condition b) of Theorem 2 is satis`ed for each pair of shared subcarriers. As an illustrative example, we consider the same ISI channel and the same transmit power constraint for all users. Then, an optimum subcarrier allocation is obtained if all users share all their subcarriers.

We conclude from part c) of our theorem that subcarriers not allocated to any user have smaller subchannel gain for some user m than any other subcarrier that is used by this speci`c user m.

We close this section by recalling that the theme of this paper relies on the availability of CSI at the transmitters. When this is impossible to acquire, the wireless multicarrier scheme of [10] that relies on linear precoding and block spreading appears most attractive. When partial CSI can be made available, then statistical water-pouring alternatives are well motivated.

V. Numerical Examples

As a `rst example to illustrate the properties studied in Theorem 2, we consider a system with two users each having unit power, and transmitting blocks of $N = 8$ symbols over ISI channels of order $L = 3$. The taps for channel 1 are $[-0.216 - 0.573j, -0.833 + 0.596j, 0.063 + 0.595j, 0.144 - 0.019j]$, and the taps for channel 2 are $[0.163 - 0.294j, 0.087 + 1.092j, -0.093 - 0.068j, 0.363 + 0.057j]$. The noise is assumed to be white with unit variance. To obtain the optimal power loadings $\lambda_{m,n}$ for each user, we have run the simpli`ed algorithm developed in Section III. We plot in Fig. 2 the optimal power spectral density (PSD) of the transmitted signals $\{\tilde{u}_m(n)\}_{m=1}^{2}$ at the frequency grid along with the subchannel gain ratios $|H_1(e^{j2\pi f})|^2/|H_2(e^{j2\pi f})|^2$, and $|H_1(e^{j2\pi f})|^2/|H_2(e^{j2\pi f})|^2$. As predicted by part a) of Theorem 2, user 1 does not allocate power at the frequencies where $|H_1(e^{j2\pi f})|^2/|H_2(e^{j2\pi f})|^2$ is relatively small unless power has been allocated at frequencies where $|H_1(e^{j2\pi f})|^2/|H_2(e^{j2\pi f})|^2$ is high. User 2 subcarriers are allocated similarly.

In a second example we consider three users with the same setup as in the `rst example, except that now our block size is $N = 12$, and the taps for channel 1 are $[0.216 - 0.11j, 0.024 - 0.107j, -0.213 - 0.276j, 0.218 + 0.099j]$, the taps for channel 2 are $[-0.241 - 0.138j, 0.042 - 0.205j, -0.491 + 0.277j, -0.186 - 0.154j]$, and the taps for channel 3 are $[-0.064 + 0.144j, -0.068 - 0.217j, -0.418 - 0.529j, -0.469 - 0.241j]$. We observe from this example that the convergence of Algorithm 1 is fast. After only $N_I = 5$ iterations the relative error $(\mathcal{C}_S - g(\{\lambda_{m,n}\}_{m=1,n=0}^{M,N-1}))/\mathcal{C}_S$ is below 10^{-4}. As before, we plot in Fig. 3 the optimal transmit-PSD along with the channel gains, namely, $|H_1(e^{j2\pi f})|^2$, $|H_2(e^{j2\pi f})|^2$, and $|H_3(e^{j2\pi f})|^2$. We remark that one subcarrier is shared by users 1 and 2, and that `ve subcarriers are not used by any user. Notice that for any user m, $m \in [1, 3]$, the channel gain at the unused frequencies is smaller than the channel gain at frequencies that are used by user m. This example illustrates well part c) of Theorem 2.

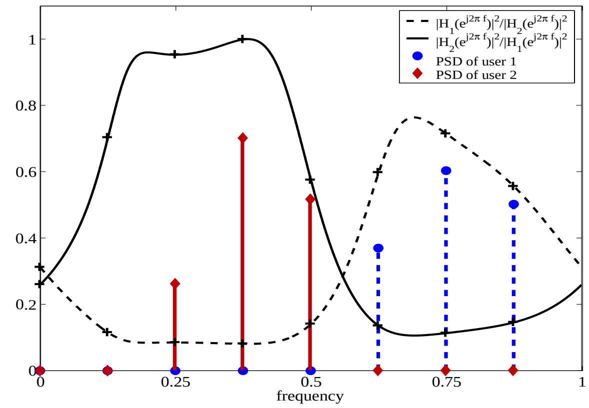

Fig. 2. Optimal transmit-PSDs and ratios of channel transfer function gains

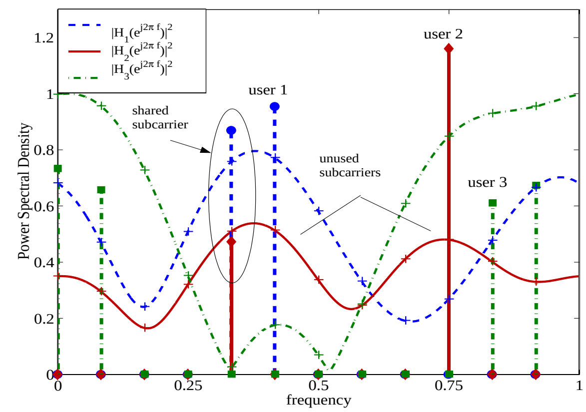

Fig. 3. Optimal transmit-PSDs along with $|H_1(e^{j2\pi f})|^2$, $|H_2(e^{j2\pi f})|^2$, and $|H_3(e^{j2\pi f})|^2$

References

[1] R. E. Blahut, *Principles and Practice of Information Theory*, Addison Wesley Longman, Inc., 1987.

[2] R. S. Cheng and S. Verdu, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-`lling," *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[4] S. Diggavi, "Multiuser DMT: a multiple access modulation scheme," in *Proc. of Globecom*, London, UK, Nov. 1996, vol. 2, pp. 1566–1570.

[5] M. C. Ferris and O. L. Mangasarian, "Parallel variable distribution," *SIAM Journal on Optimization*, vol. 4, pp. 815–832, 1994.

[6] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, Inc., 1968.

[7] S. Kasturia, J. T. Aslanis, and J. M. Ciof`, "Vector coding for partial response channels," *IEEE Trans. on Info. Theory*, pp. 741–761, 1990.

[8] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for multiple access through ISI channels," *IEEE Trans. on Information Theory*, June 2001, submitted; see also *Proc. of 39th Allerton Conf.*, Monticello, IL, Oct. 3-5, 2001.

[9] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. on Info. Theory*, pp. 1019–1032, April 1999.

[10] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[11] W. Yu, W. Rhee, S. Boyd, and J. M. Ciof`, "Iterative water-`lling for gaussian vector multiple access channel," *IEEE Intl. Symp. on Info. Theory*, p. 322, June 2001.

[12] C. Zeng, L. Hoo, and J. M. Ciof`, "Ef`cient water-`lling algorithm for a Gaussian multiaccess channel with ISI," in *IEEE Vehicular Techn. Conf.*, Boston, MA, Sept. 2000, pp. 1072–1077.